Journal of Alasmarya University: Basic and Applied Sciences
Volume (5) Issue 1 (June 2020)

مجلة الجامعة الأسمرية: العلوم الأساسية والتطبيقية

المجلد (5) العدد 1 (يونيو 2020)

A NEW MODIFICATION INTO SELF-SCALING BFGS VARIABLE METRIC METHOD

[*]Basim A. Hassan, and Ussein O. Hhalo

Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul, IRAQ

[*]basimah@uomosul.edu.iq

Abstract
In this paper, we propose a modified version of the self-scaling Yuan's [14] update which is based on the simple idea of approximation the objective function by technique is induced. Arithmetical Performance signifies that the new proposed techniques are more well-organized than the ordinary BFGS-technique.

Keywords: Variable metric method, Self-scaling variable metric method, convergence properties.

1. Introduction :

Variable metric methods are a broadly utilized category from iterative techniques for solving the problem :

$$\min\{f(x) \mid x \in R^n\}. \qquad \text{..........(1)}$$

They are iterative. On the kth iteration of the variable metric methods, a symmetric and nonnegative definite matrix B_k is certain, and a search direction is calculated by :

$$d_k = -B_k^{-1} g_k, \qquad ....(2)$$

where g_k is the gradient of function f. One then computes the new iterate by :

$$x_{k+1} = x_k + \alpha_k d_k \qquad ....(3)$$

where the step length α_k satisfies the Wolfe conditioned :

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta_1 \alpha_k d_k^T g_k \qquad ....(4)$$

$$g(x_k + \alpha_k d_k)^T d_k \ge \delta_2 d_k^T g_k \qquad ....(5)$$

where $0 < \delta_1 \le \delta_2 < 1$. More details can be found in Fletcher [7,12].

The selection matrices B_k from necessary options that need B_k positive definite and gratifying the quasi-Newton equation :

$$B_{k+1} s_k = y_k \qquad ....(6)$$

where

$$s_k = x_{k+1} - x_k = \alpha_k d_k \text{ and } y_k = g_{k+1} - g_k \qquad ....(7)$$

The step size is calculated by :

$$\alpha_k = -\frac{g_k^T d_k}{d_k^T G d_k}. \qquad ....(8)$$

The direction d_k in (2) is that the resolution of the subsequent quadratic sub drawback :

$$f_{k+1} = f_k + d_k^T g_k + \frac{1}{2} d_k^T B_k d_k . \qquad ....(9)$$

which is an estimate to problem (1) next to the recent iterate x_k, for small d_k.

One well-known update formula is the BFGS formula, which updates B_{k+1} from B_k, y_k and s_k in the following way :

$$B_{k+1}^{BFGS} = B_k - \frac{B_k s_k s_k^T B_k^{\ T}}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k} \qquad ....(10)$$

Let H_k be the inverse of B_k. Then the inverse update formula of (9) the method is represented as :

$$H_{k+1}^{BFGS} = H_k - \frac{H_k y_k s_k^T + s_k y_k^T H_k}{s_k^T y_k} + \left[1 + \frac{y_k^T H_k y_k}{s_k^T y_k}\right] \frac{s_k s_k^T}{s_k^T y_k} \qquad ....(11)$$

More details can be found in Ladislav [9]. For more studies and recent references on the variable metric, the interested researcher may refer to [1-5].

Below, I will now modify the derivation of Yuan's self-scaling method and study its convergence.

2. New Modification into Yuan's Self-scaling Variable Metric

Based on the idea of Yuan's, we present a new Modification self-scaling variable metric.

The idea of Yuan's [14], important idea to derive the modified BFGS update as follows :

$$B_{k+1}^{BFGS} = B_k - \frac{B_k s_k s_k^T B_k{}^T}{s_k^T B_k s_k} + \rho_k \frac{y_k y_k^T}{s_k^T y_k} \qquad \text{..........(12)}$$

where

$$\rho_k = 2\left[f_k - f_{k+1} + g_{k+1}^T s_k\right] / y_k^T s_k . \qquad \text{..........(13)}$$

The equation (12) can be rewritten as :

$$s_k^T B_{k+1} s_k = 2\left[f_k - f_{k+1} + g_{k+1}^T s_k\right]. \qquad \text{..........(14)}$$

Note that the updates (12) and the usual BFGS updates are very different.

In fact, equation (13) with an exact line search $g_{k+1}^T d_k = 0$ can be rearranged as the form :

$$s_k^T B_{k+1} s_k = 2\left[f_k - f_{k+1}\right]. \qquad \text{..........(15)}$$

The above equation can be rewritten as :

$$\rho_k^{BK1} = 2\left[f_k - f_{k+1}\right] / y_k^T s_k . \qquad \text{..........(16)}$$

However, condition (15) may be modified further to give :

$$\begin{aligned} s_k^T B_{k+1} s_k &= f(x_k) - f(x_{k+1}) - \alpha_k g_k^T d_k / 2 \\ s_k^T B_{k+1} s_k &= f(x_k) - f(x_{k+1}) - g_k^T s_k / 2 \end{aligned} \qquad \text{..........(17)}$$

Now, equation (17) can be rewritten as :

$$\rho_k^{BK2} = \left[f(x_k) - f(x_{k+1}) - g_k^T s_k / 2\right] / y_k^T s_k . \qquad \text{..........(18)}$$

He showed that a modified QN-algorithm could be written as follows :

$$H_{k+1} = H_k - \frac{H_k y_k s_k^T + s_k y_k^T H_k}{s_k^T y_k} + \left[\frac{1}{\rho_k} + \frac{y_k^T H_k y_k}{s_k^T y_k} \right] \frac{s_k s_k^T}{s_k^T y_k} \quad . \qquad \text{..........(19)}$$

With this modification, we present the following algorithm.

New algorithm:

1 : Given $x_0 \in R^n$, $H_0 = I$ and $k = 0$.

2 : Calculate g_k, if $\|g_k\| < \varepsilon$ then stop.

3 : Generate $d_k = -H_k g_k$.

4 : Calculate α_k such that (4)-(5) hold.

5 : Let the next iterative be $x_{k+1} = x_k + \alpha_k d_k$.

6 : Compute ρ_k, and update H_{k+1} by (8) and (18).

7 : $k = k+1$ and go to step 2.

3. Global convergence analysis

We reading the convergence of the new methods with the normal states is needed. The Hessian matrix off is indicated by G.

Assumptions 3.1.

(1) The target perform f is convex, finite below, and doubly unendingly differentiable in R^n and there exists a constant f^* such that $\lim_{N \to \infty} f(x_k) = f^*$.

(2) The Hessian matrix $G(x)$ is bounded above in norm for all $x \in D = \{x \in R^n : f(x) \le f(x_1)\}$. More details can be found in [8].

In [10], we were defining the angle between s_k and $B_k s_k$, θ_k by :

$$\cos\theta_k = s_k^T B_k s_k / \|s_k\| \|B_k s_k\| = -s_k^T g_k / \|s_k\| \|g_k\| \qquad \text{..........(20)}$$

and the Rayleigh quotient $q_k = s_k^T B_k s_k / s_k^T s_k$.

Theorem 3.1

Any iterative procedure of the form (3), where α_k satisfies the states $(4-5)$. The gradient $\nabla f(x_k) = g_k$ is Lipschitz continuous on N, that is, there contain a constant $L > 0$ such that :

$$\|g(x) - g(y)\| \le L\,\|x - y\| \ ; \quad \forall\ x,\ y \in N \ . \qquad \text{..........(21)}$$

Then :

$$\sum_{k\ge 1} \|g_k\|^2 \ \cos^2\theta_k \ < \infty \ . \qquad \text{..........(22)}$$

Proof :

Using (4) and (20) we obtain :

$$f(x_k + \alpha_k d_k) \le f(x_k) - c\,\frac{(s_k^T g_k)^2}{\|s_k\|} \le f(x_k) - c\|g_k\|^2 \cos^2\theta_k \quad . \qquad \text{..........(23)}$$

with some constant $c > 0$. On the other hand, from (1), we have :

$$\begin{aligned}\sum_{k=1}^{\infty}\big(f(x_k) - f(x_{k+1})\big) &= \lim_{N\to\infty}\sum_{k=1}^{N}\big(f(x_k) - f(x_{k+1})\big)\\ &= \lim_{N\to\infty}\big(f(x_1) - f(x_{k+1})\big) \qquad \text{..........(24)}\\ &= f(x_1) - f^*\end{aligned}$$

which combining with the WWP rule (4) yields :

$$\sum_{k\geq 1}^{\infty}\|g_k\|^2 \cos^2\theta_k < \infty . \qquad \text{..........(25)}$$

which concludes the proof.

4. Arithmetical Performance

In this section, arithmetical Performance is reported. We tested the variable metric algorithms with the following ρ_k. We used test problems from Mor´e, Garbow, and Hillstrom [11]. They are summarized in Table 1. The line search calculates α_k gratifying the states (5-6) with $\delta_1 = 0.001$ and $\delta_2 = 0.9$. The stopping condition: " if $|f(x_k)| > 10^{-5}$, let $stop1 = |f(x_k) - f(x_{k+1})| \,/\, |f(x_k)|$; Otherwise, let $stop1 = |f(x_k) - f(x_{k+1})|$. For each problem, if $\|g_k\| < \varepsilon$ or $stop1 < 10^{-5}$ was satisfied, the program will be stopped. The following Himmeblau stop rule is used [13]. The program was also terminated if the number of iterations exceeds 1000".

We use MATLAB to check the elected problems. The arithmetical Performance are listed in Table 1: where NI and NF stand for the total number of all iterations and the total number of function evaluations, respectively. The results printed in boldface imply that algorithm New methods performed better than algorithm BFGS comparing in each column.

Table 1 : Comparison of different BFGS-algorithms with different test functions and different dimensions

		BFGS algorithm		BFGS with ρ_k^{BK1}		BFGS with ρ_k^{BK2}	
P. No.	**n**	**NI**	**NF**	**NI**	**NF**	**NI**	**NF**
Rose	2	35	140	7	90	41	142
Froth	2	9	26	3	31	3	31
Badscp	2	43	166	4	34	4	34
Badscb	2	3	30	3	30	3	30
Beale	2	15	50	4	35	6	43
Jensam	2	2	27	2	27	2	27
Helix	3	34	113	37	111	34	107

Bard	3	16	54	16	52	15	49
Gauss	3	2	4	2	4	2	4
Gulf	3	2	27	2	27	2	27
Box	3	2	27	2	27	2	27
Sing	4	20	60	25	82	26	86
Wood	4	19	61	3	31	3	31
Kowosb	4	21	65	19	58	17	77
Bd	4	17	54	3	31	3	31
Osb1	5	2	27	2	27	2	27
Biggs	6	25	72	F	F	59	238
Osb2	11	3	31	3	31	3	31
Watson	20	31	102	40	126	38	119
Singx	400	64	209	3	31	3	31
Pen1	400	2	27	2	27	2	27
Pen2	200	2	5	2	5	2	5
Vardim	100	2	27	2	27	2	27
Trig	500	9	33	29	99	29	99
Bv	500	2	4	2	4	2	4
Ie	500	6	61	7	19	7	19
Band	500	57	281	3	31	3	31
Lin	500	2	4	2	4	2	4
Lin1	500	3	7	3	7	3	7
Lino	500	3	7	3	7	3	7
Total		**428**	**1684**	**235**	**1115**	**264**	**1184**

The graphs are plotted using data derived from numerical computations using the output model proposed by Dolan and More [6]. The suggested BK1 and BK2 approach has the best results in terms of both number of iterations as seen in Fig (1) and number of function evaluation Fig (2).

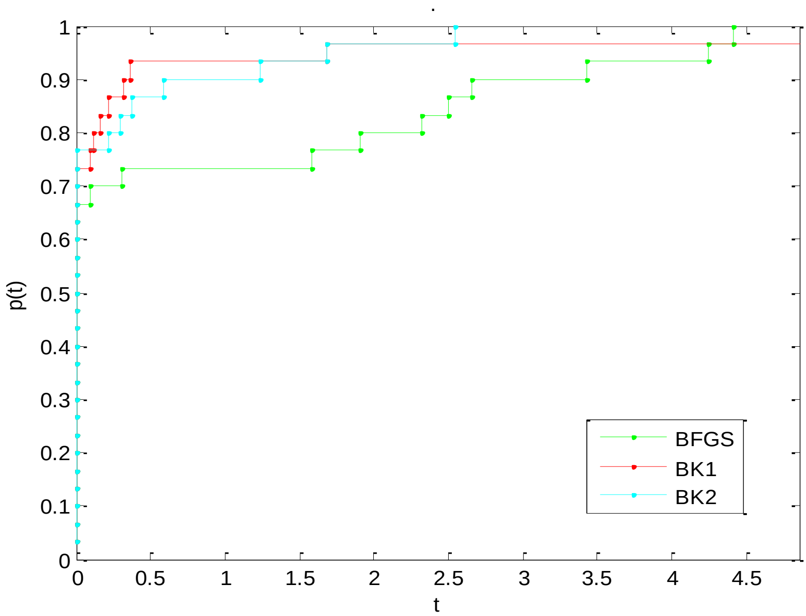
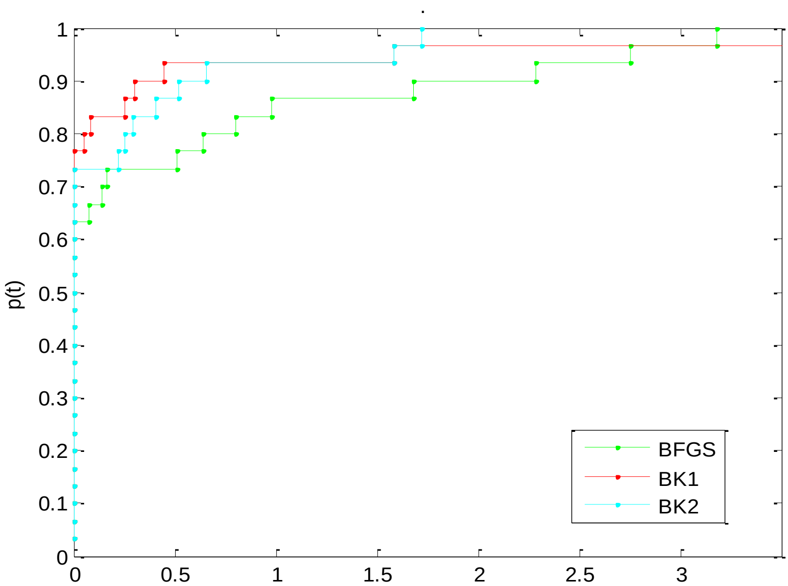


Fig (1). Number of iteration profile via Dolan and More



Fig(2). Number of function evaluation profile via Dolan and More

5. Conclusions

In this paper, we have derived a new self- scaling variable metric following Yuan's Self-scaling [14]. This is a modification of the Yuan's formula. We find that this method performed better than the original BFGS method in each table, especially in Tables 1 with has been shown to be globally convergent.

References

[1] Basim A. Hassan (2020)," A new type of quasi-Newton updating formulas based on the new quasi-Newton equation", SIAM J. N. A., control and optimization. Numerical Algebra, Control & Optimization.2, pp.227–235.

[2] Basim A. Hassan and Hawraz N. Jabbar, (2018) " A New Transformed Biggs 's Self-Scaling Quasi-Newton Method for Optimization", ZANCO Journal of Pure and Applied Sciences, 31, pp. 1-5.

[3] Basim A. Hassan and Ghada M., (2020)," A New Quasi-Newton Equation on the Gradient Methods for Optimization Minimization Problem", Indonesian Journal of Electrical Engineering and Computer Science, 19,(2), pp. 737–744.

[4] Basim A. Hassan and Mohammed W. T., (2020)," A New Variants of Quasi-Newton Equation Based on the Quadratic Function for Unconstrained Optimization", Indonesian Journal of Electrical Engineering and Computer Science, 19,(2), pp.701–708.

[5] Basim A. Hassan and Mohammed W. T, (2019)."A Modified Quasi-Newton Equation in the Quasi-Newton Methods for Optimization", HIKARI Ltd, Applied Mathematical Sciences, 10, pp. 463 – 472.

[6] Dolan E. and. Moré J. J, '(2002), Benchmarking optimization software with performance profiles. Mathematical Programming. 91(2),pp.201-213.

[7] Fletcher, R. (1987), "Practical Methods of Optimization" .John Wiley and Sons, Chi Chester (New York).

[8] Nocedal, J. and Wright J. (2006), "Numerical Optimization", Springer Series in Operations Research, Springer Verlag, New York, USA.

[9] Ladislav L., (1982), "Variable metric method with limited storage for large-scale unconstrained minimization", Kybernetika J., 18, pp.517-528.

[10] Liao A., (1997)," Modifying the BFGS method"', Operations Research Letters, 20, pp. 171-177.

[11] More J., Garbow B., and Hillstrome K., (1981)," Testing unconstrained optimization software", ACM Trans. Math. Software, 7,pp. 17-41.

[12] Wolfe P.,(1971)," Convergence conditions for ascent methods", (II): some corrections. SIAM Review, 13, pp.185-188.

[13] Yuan Y. and Sun W., (1999)," Theory and Methods of Optimization", Science Press of China.

[14] Yuan, Y. (1991), "A modified BFGS algorithm for unconstrained optimization". IMA Journal Numerical Analysis, 11, pp. 325-332.